UCLA TECHNICAL REPORT
20 JUNE 1977

PUPILLOMETRIC MEASUREMENT OF COGNITIVE WORKLOAD IN PROCESSING NUMERICAL VISUAL DISPLAYS

JACKSON BEATTY
DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF CALIFORNIA, LOS ANGELES

THIS RESEARCH WAS SUPPORTED BY THE ADVANCED RESEARCH PROJECTS AGENCY OF THE DEPARTMENT OF DEFENSE AND WAS MONITORED BY ONR UNDER CONTRACT NO. N00014-70-C-0350.

Summary

The momentary workload that is imposed by a cognitive task upon the limited capacity human information-processing system appears to be accurately reflected in the momentary level of central nervous system activation. The utility of pupillometric methods of workload assessment is tested in an experiment involving the perception of four-digit visual displays of varying degrees of clarity. Pupillary dilations accurately reflect the cognitive load imposed by these displays, even in the absence of encoding errors.

A

INTRODUCTION

Information processing tasks differ in the extent and duration of the demands that they place upon the limited capacity of the human nervous system to handle information. For most tasks, processing demands are not constant, but vary from moment to moment in response to changes in the functional organization of the task. These demands may be thought to represent the cognitive workload associated with the task, a time-varying function of the demand for limited resources.

Given the assumption that cognitive capacity is fixed (Broadbent, 1958), the momentary demands of any single processing function for capacity may be estimated by determining the amount of residual capacity that may be allocated to another processing task that is assigned a secondary priority (Kerr, 1973). Secondary-task measurement of cognitive workload is of major importance in the study of both cognitive capacity and the resource demands of particular cognitive processes, but both technical (Kerr, 1973) and theoretical (Norman & Bobrow, 1975) difficulties preclude the utilization of secondary-task procedures in many situations. For this reason the more convenient method of subjective estimation of cognitive workload is still commonly employed (McCormick, 1970) despite serious questions as to both the reliability and validity of such rating procedures.

A third approach to the problem of measuring momentary cognitive workload stems from the observation that momentary workload is directly reflected in the momentary level of central nervous sytem (CNS) activation (Kahneman, 1973; Pribram & McGuinness, 1975). Of the various indicators of activation, pupillometric measurement techniques (Lowenfeld, 1958; Hess & Polt, 1964; Goldwater, 1972) appear to be most sensitive and reliable (Kahneman, Tursky, Shapiro, & Crider, 1969).

Perceptual processes appear to proceed quite effortlessly and place rather little demand upon the limited capacity of the human information-processing system (Kahneman, 1973). Thus Wickens (1974) was unable to observe a secondary task decrement when a sensory signal-detection task was imposed as the primary task in an experiment investigating the distribution of processing capacity. The workload involved in the detection of weak signals is quite small.

In this context, it is of interest to note that small but reliable pupillary dilations accompany the detection of both visual and acoustic signals at near-threshold intensities. Hakerem and Sutton (1966) examined the pupillary movements that accompany the perception of weak visual stimuli and were able to show a dilation for signals that were detected which was absent for signals that were missed. More recently Beatty and Wagoner (1975) provided a pupillometric analysis of activation in the detection of weak acoustic signals using a rating-scale response procedure (see Green & Swets, 1966). Using unmarked observation intervals, no pupillary dilations were observed in the absence of a signal regardless of the outcome of the observer's decision. In the presence of a signal, a dilation of the pupil appeared in the interval between signal delivery and response cue onset. The magnitude of this dilation varied monotonically with the observer's rated probability that a signal had been presented.

These data raise the interesting possibility that pupillometric methods may provide a more sensitive measure of cognitive load than do conventional secondary-task measurement techniques. Thus the small pupillary dilations observed during perceptual processing may be indexing brain workload levels that are not of sufficient magnitude to be detected by secondary task interference methods.

To test the hypothesis that pupillometric measures might reflect the cognitive workload involved in encoding displays, pupillary measures were obtained as subjects observed a four-digit computer-controlled visual display at three levels of visual noise, which is known to selectively affect the encoding state in perceptual processing (Sternberg, 1969).

METHOD

Subjects: Twelve undergraduate students and laboratory personnel served as volunteer subjects in this experiment. The students were paid $2.50 for their services.

Procedure: Each subject was tested individually in a sound-attenuated experimental chamber, containing the pupillometric computer terminal. This interface consists of a computer-controlled cathode ray tube (CRT) display, a microswitch keyboard, and a television camera with headrest which provides an image of the eye for pupillometric measurement. Room illumination was 17 foot-lamberts.

The task of the subject was to read a four-digit number displayed on the computer-controlled CRT and to report the digit string at the end of each trial. The illumination of the display was approximately constant through all phases of the experiment. Between trials, the CRT displayed four 10 by 16 fields of regularly-spaced points. At the onset of a trial, initiated by a button press by the subject, these arrays were replaced by four randomly arranged arrays for 500 msec, following which four digits were presented, again for 500 msec. The four random fields then returned for the remainder of the trial (2 sec.). The trial was terminated by four downward-pointing arrows, indicating to the subject that he should enter the number on the keyboard.

On the average, all display fields contained the same number of points, insuring that the luminance of the display did not change during a trial.

The experiment was composed of 60 trials, 20 of which presented stimuli degraded at one of three levels of visual noise (.1, .2, or .3). The ordering of trials was completely randomized for each subject.

The level of visual noise was defined by the probability that a point used in generating a character would be displaced to a randomly selected position in the background area of the 10 by 16 display space for that character. Thus the character display becomes increasingly difficult to encode as the level of visual noise is increased.

Data analysis. Pupillometric measurements were made every 25 msec during each trial of the experiment and stored on disc for later verification and analysis. Verification was accomplished by visual inspection of individual evoked pupillary responses. Verification was performed blindly with respect to stimulus type and response correctness. All trials containing major artifacts were discarded. Trials with small artifacts during non-critical periods were corrected by linear interpolation. The resulting artifact-free data were then averaged separately for each of the three levels of visual noise over all trials in which the subject responded without error.

The dilation associated with the correct processing of the numerical stimuli at each of the three levels of visual degradation was then estimated by computing the difference in pupillary diameter of the averaged pupillary response at display onset and in the .5 sec interval preceding the onset of the response cue. These values were then statistically analyzed using the analysis of variance (ANOVA) for repeated measures.

The percent errors was also computed for each subject and each stimulus condition to obtain a behavioral measure of task difficulty.

RESULTS

Pupillary dilation between the onset of the stimulus and the onset of the response cue provides an accurate index of encoding difficulty for the four-digit visual display. Figure 1 presents the averaged evoked pupillary responses for correctly reported visual displays as a function of visual noise level. Pre-response pupillary dilation appears to vary as a direct function of visual noise level. The results of an ANOVA on these data confirm this observation. The effect of visual noise level is highly significant in these data ($F[2,22] - 10.991$; $p < .01$).

Mean pupillary dilation as a function of visual noise level is shown in Figure 2, as is the percent errors for each stimulus condition. Like the pupillometric measure, the behavioral indicator of encoding difficulty, response accuracy, also varies significantly across stimulus conditions ($F[2,22] = 46.679$; $p < .001$). The similarity of the two functions is evident in Figure 2. It should be remembered, however, that the pupillometric function is computed entirely from trials in which the subject responded correctly. Thus it appears that greater activation as measured pupillometrically is associated with correctly encoding visually degraded information and that the greater processing difficulty associated with the higher levels of visual noise is confirmed by the performance data.

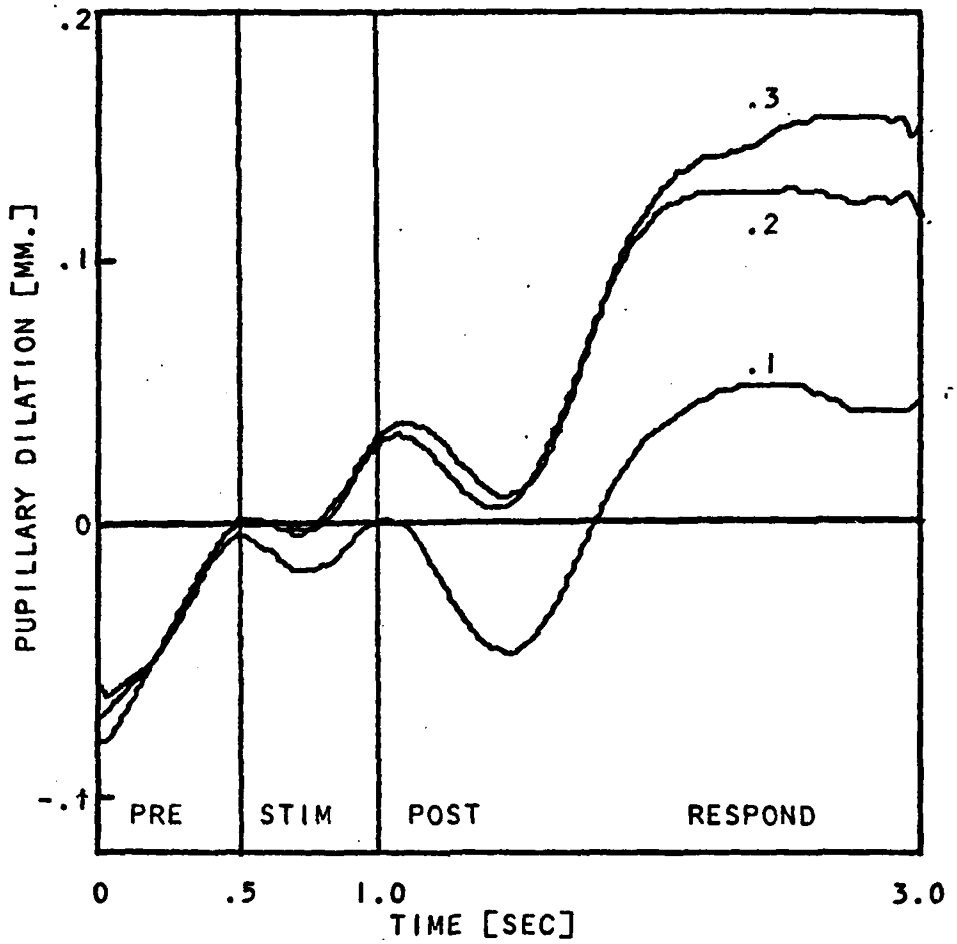

Figure 1. Averaged evoked pupillary responses for all subjects as a function of visual noise level. Larger processing dilations are associated with encoding the more severely degraded visual displays.

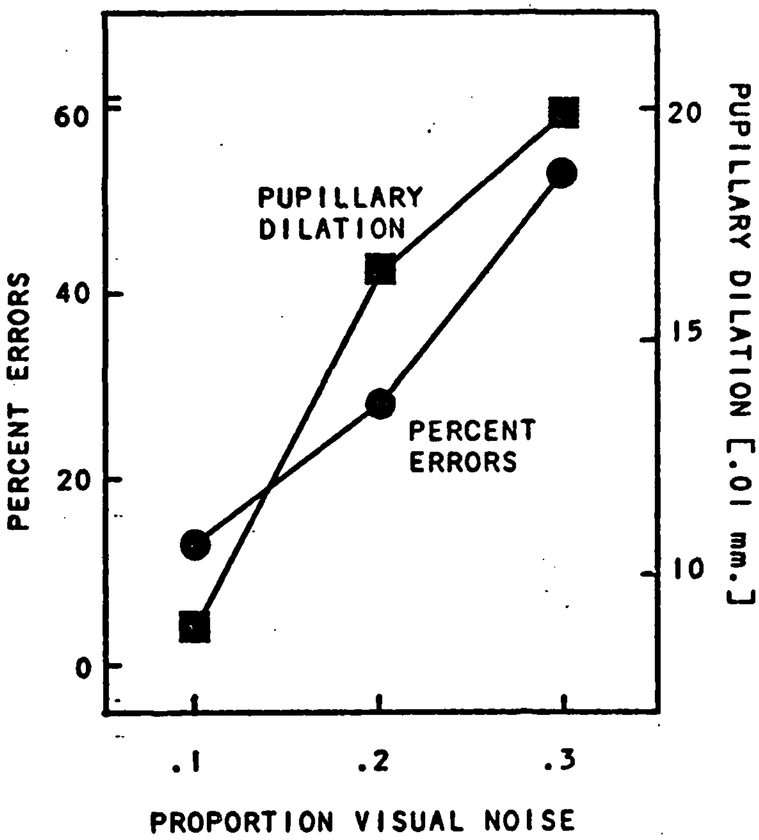

Figure 2. Mean pupillary dilation and percent errors as a function of level of visual noise. Pupillary dilation was computed from error-free trials.

DISCUSSION

Traditional interference and subjective-rating methods of workload evaluation have been employed in the design of complex man/machine interfaces, but neither is without its own particular limitations. Pupillometric methods of workload estimation provide a third alternative that in certain situations might be preferable to either of the more traditional measurements.

One problem for which pupillometric assessment procedures appear to be well-suited is that of display evaluation. Pupillometric methods permit reliable measurement of the small cognitive workloads associated with the processing of sensory information that may not be detectable by interference methods. The present experiment confirms the effects of display readability on the pupillary dilations accompanying information acquisition.

The most intriguing possibility is that the measurement of central nervous system activation associated with cognitive function might provide a common metric for the comparison of workload in tasks that differ substantially in their functional characteristics. Underlying this possibility is the idea that CNS activation is the limited general resource that is allocated among cognitive processes demanding capacity. If this is the case, then it may be possible to directly compare perceptual, memory, symbol manipulation and response processes in terms of activation requirements. At present, we may safely conclude that the pupillometric measures of activation are useful in measuring cognitive load for a range of cognitive processes, including the encoding of visual information.

REFERENCES

Beatty, J.; and Wagoner, B. L.: Pupillary Measurement of Sensory and Decision Processes in an Acoustic Signal-detection Task. Paper presented at the meeting of the Psychonomic Society, Denver, Colorado, 1975.

Broadbent, D. E.: Perception and Communication. London: Pergamon Press, 1958.

Goldwater, B. C.: Psychological Significance of Pupillary Movements. Psychological Bulletin, 1972, 77, 340-355.

Green, D. M.; and Swets, J. A.: Signal Detection Theory and Psychophysics. New York: Wiley, 1966.

Hakerem, G.; and Sutton, S.: Pupillary Response at Visual Threshold. Nature, 1966, 212, 485.

Hess, E. H.; and Polt, J. H.: Pupil Size in Relation to Mental Activity During Simple Problem Solving. Science, 1964, 143, 1190-1192.

Kahneman, D.: Attention and Effort. Englewood Cliffs, New Jersey: Prentice-Hall, 1973.

Kahneman, D.; Tursky, B.; Shapiro, D.; and Crider, A.: Pupillary, Heart Rate and Skin Resistance Changes During a Mental Task. Journal of Experimental Psychology, 1969, 79, 164-167.

Kerr, B.: Processing Demands During Mental Operations. Memory and Cognition, 1973, 1, 401-412.

Loewenfeld, I. E.: Mechanisms of Reflex Dilations of the Pupil, Historical and Experimental Analysis. Documents Ophthalmologica, 1959, 12, 185-448.

McCormick, E. J.: Human Factors Engineering. (3rd ed.) New York: McGraw-Hill, 1970.

Norman, D. A.; and Bobrow, D. G.: On Data-limited and Resource-limited Processes. _Cognitive Psychology_, 1975, _7_, 44-64.

Pribram, K. H.; and McGuinness, D.: Arousal, Activation, and Effort in the Control of Attention. _Psychological Review_, 1975, _82_, 116-149.

Wickens, C. D.: The Effect of Time Sharing on the Performance of Information Processing Tasks: A Feedback Control Analysis. Technical Report No. 51. Human Performance Center, the University of Michigan, 1974.